O ) 10
/0 2 4

# RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: ___10 028, 075 A___
Source: ___O1PE___
Date Processed by STIC: ___10 30 02___

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
 PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
 PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
 **http://www.uspto.gov/web/offices/pac/checker**

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<**http://www.uspto.gov/ebc/efs/downloads/documents.htm>** , EFS Submission User Manual - ePAVE)
2. **U.S. Postal Service:** U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
         Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

O IPE

| ERROR DETECTED | SUGGESTED CORRECTION | SERIAL NUMBER: 10 028,075A |
|---|---|---|

**ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE**

1_____Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length

Sequence(s)_____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. **This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.**

7_____Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)      SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10_____Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial-Sequence

11_√_Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13_____Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

OIPE

Does Not Comply
Corrected Diskette Needed

RAW SEQUENCE LISTING          DATE: 10/30/2002
PATENT APPLICATION:  US/10/028,075A          TIME: 15:01:56

Input Set : **A:\EP.txt**
Output Set: **N:\CRF4\10302002\J028075A.raw**

```
 3 <110> APPLICANT: Khan, Nisar A.
 4         Benner, Robert
 6 <120> TITLE OF INVENTION: Gene regulator
 8 <130> FILE REFERENCE: 2183-5223US
10 <140> CURRENT APPLICATION NUMBER: 10/028,075A
11 <141> CURRENT FILING DATE: 2001-12-21
13 <150> PRIOR APPLICATION NUMBER: EP 01203748.7
14 <151> PRIOR FILING DATE: 2001-10-04
16 <160> NUMBER OF SEQ ID NOS: 175
18 <170> SOFTWARE: PatentIn Ver. 2.1
20 <210> SEQ ID NO: 1
21 <211> LENGTH: 4
22 <212> TYPE: PRT
23 <213> ORGANISM: Artificial Sequence
25 <220> FEATURE:
26 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
28 <400> SEQUENCE: 1
29 Leu Gln Gly Val
30    1
32 <210> SEQ ID NO: 2
33 <211> LENGTH: 4
34 <212> TYPE: PRT
35 <213> ORGANISM: Artificial Sequence
37 <220> FEATURE:
38 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
40 <400> SEQUENCE: 2
41 Ala Gln Gly Val
42    1
43 <210> SEQ ID NO: 3
44 <211> LENGTH: 6
45 <212> TYPE: PRT
46 <213> ORGANISM: Artificial Sequence
48 <220> FEATURE:
49 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
51 <400> SEQUENCE: 3
52 Val Leu Pro Ala Leu Pro
54    1                   5
56 <210> SEQ ID NO: 4
57 <211> LENGTH: 16
58 <212> TYPE: PRT
59 <213> ORGANISM: Artificial Sequence
61 <220> FEATURE:
62 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide
```

*The type of errors shown exist throughout the Sequence Listing. Please check subsequent sequences for similar errors.*

— Must explain Genetic source —
See error summary sheet item 11

```
 64 <400> SEQUENCE: 4
 65 Met Leu Ala Arg Arg Lys Pro Val Leu Pro Ala Leu Thr Ile Asn Pro
 66   1               5                  10                 15
 68 <210> SEQ ID NO: 5
 69 <211> LENGTH: 7
 70 <212> TYPE: PRT
 71 <213> ORGANISM: Artificial Sequence
 73 <220> FEATURE:
 74 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide
 76 <400> SEQUENCE: 5
 77 Met Leu Ala Arg Arg Lys Pro
 78   1               5
 80 <210> SEQ ID NO: 6
 81 <211> LENGTH: 4
 82 <212> TYPE: PRT
 83 <213> ORGANISM: Artificial Sequence
 85 <220> FEATURE:
 86 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide
 88 <400> SEQUENCE: 6
 89 Met Leu Ala Arg
 90   1
 92 <210> SEQ ID NO: 7
 93 <211> LENGTH: 6
 94 <212> TYPE: PRT
 95 <213> ORGANISM: Artificial Sequence
 97 <220> FEATURE:
 98 <223> OTHER INFORMATION: Description of Artificial Sequence: peptide
100 <400> SEQUENCE: 7
101 Val Leu Pro Ala Leu Thr
102   1               5
104 <210> SEQ ID NO: 8
105 <211> LENGTH: 5
106 <212> TYPE: PRT
108 <213> ORGANISM: Artificial Sequence
111 <220> FEATURE:
112 <223> OTHER INFORMATION: Description of Artificial Sequence:
113       pdb/1QMH/1QMH-A
115 <400> SEQUENCE: 8
116 Val Leu Pro Ala Leu
117   1               5
119 <210> SEQ ID NO: 9
120 <211> LENGTH: 4
121 <212> TYPE: PRT
122 <213> ORGANISM: Artificial Sequence
124 <220> FEATURE:
125 <223> OTHER INFORMATION: Description of Artificial Sequence:
126       pdb/4NOS/4NOS-A
128 <400> SEQUENCE: 9
129 Phe Pro Gly Cys
```

RAW SEQUENCE LISTING                    DATE: 10/30/2002
PATENT APPLICATION:  US/10/028,075A     TIME: 15:01:56

Input Set : A:\EP.txt
Output Set: N:\CRF4\10302002\J028075A.raw

```
130   1
132 <210> SEQ ID NO: 10
133 <211> LENGTH: 4
134 <212> TYPE: PRT
135 <213> ORGANISM: Artificial Sequence
137 <220> FEATURE:
138 <223> OTHER INFORMATION: Description of Artificial Sequence: Hs.297775.1
140 <400> SEQUENCE: 10
141 Pro Gly Cys Pro
142   1
144 <210> SEQ ID NO: 11
145 <211> LENGTH: 7
146 <212> TYPE: PRT
147 <213> ORGANISM: Artificial Sequence
149 <220> FEATURE:
150 <223> OTHER INFORMATION: Description of Artificial Sequence:
151       swiss/P81272/NS2B HUMAN
153 <400> SEQUENCE: 11
154 Gly Val Leu Pro Ala Val Pro
155   1               5
157 <210> SEQ ID NO: 12
158 <211> LENGTH: 6
159 <212> TYPE: PRT
160 <213> ORGANISM: Artificial Sequence
162 <220> FEATURE:
163 <223> OTHER INFORMATION: Description of Artificial Sequence:
164       swiss/P81272/NS2B HUMAN
166 <400> SEQUENCE: 12
167 Val Leu Pro Ala Val Pro
168   1               5
170 <210> SEQ ID NO: 13
171 <211> LENGTH: 4
172 <212> TYPE: PRT
173 <213> ORGANISM: Artificial Sequence
175 <220> FEATURE:
176 <223> OTHER INFORMATION: Description of Artificial Sequence:
177       pdb/1FZV/1FZV-A
179 <400> SEQUENCE: 13
180 Pro Ala Val Pro
181   1
183 <210> SEQ ID NO: 14
184 <211> LENGTH: 9
185 <212> TYPE: PRT
186 <213> ORGANISM: Artificial Sequence
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
191 <400> SEQUENCE: 14
192 Leu Gln Gly Val Val Pro Arg Gly Val
193   1               5
```

```
195 <210> SEQ ID NO: 15
196 <211> LENGTH: 4
197 <212> TYPE: PRT
198 <213> ORGANISM: Artificial Sequence
200 <220> FEATURE:
201 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
203 <400> SEQUENCE: 15
204 Gly Val Val Pro
205   1
207 <210> SEQ ID NO: 16
208 <211> LENGTH: 5
209 <212> TYPE: PRT
210 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
213 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
216 <400> SEQUENCE: 16
217 Val Pro Arg Gly Val
218   1                 5
222 <210> SEQ ID NO: 17
223 <211> LENGTH: 4
224 <212> TYPE: PRT
225 <213> ORGANISM: Artificial Sequence
227 <220> FEATURE:
228 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
230 <400> SEQUENCE: 17
231 Pro Arg Gly Val
232   1
234 <210> SEQ ID NO: 18
235 <211> LENGTH: 5
236 <212> TYPE: PRT
237 <213> ORGANISM: Artificial Sequence
239 <220> FEATURE:
240 <223> OTHER INFORMATION: Description of Artificial Sequence: polypeptide
242 <400> SEQUENCE: 18
243 Met Ala Pro Lys Lys
244   1
246 <210> SEQ ID NO: 19
247 <211> LENGTH: 4
248 <212> TYPE: PRT
249 <213> ORGANISM: Artificial Sequence
251 <220> FEATURE:
252 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
254 <400> SEQUENCE: 19
255 Leu Gln Gly Ala
256   1
258 <210> SEQ ID NO: 20
259 <211> LENGTH: 10
260 <212> TYPE: PRT
261 <213> ORGANISM: Artificial Sequence
```

```
                RAW SEQUENCE LISTING               DATE: 10/30/2002
                PATENT APPLICATION:  US/10/028,075A   TIME: 15:01:56

                Input Set : A:\EP.txt
                Output Set: N:\CRF4\10302002\J028075A.raw
```

```
263 <220> FEATURE:
264 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
266 <400> SEQUENCE: 20
267 Val Leu Pro Ala Leu Pro Gln Val Val Cys
268   1              5              10
270 <210> SEQ ID NO: 21
271 <211> LENGTH: 6
272 <212> TYPE: PRT
273 <213> ORGANISM: Artificial Sequence
276 <220> FEATURE:
277 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
279 <400> SEQUENCE: 21
280 Ala Leu Pro Ala Leu Pro
281   1              5
283 <210> SEQ ID NO: 22
284 <211> LENGTH: 6
285 <212> TYPE: PRT
286 <213> ORGANISM: Artificial Sequence
288 <220> FEATURE:
289 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
291 <400> SEQUENCE: 22
292 Val Ala Pro Ala Leu Pro
293   1           5
295 <210> SEQ ID NO: 23
296 <211> LENGTH: 7
297 <212> TYPE: PRT
298 <213> ORGANISM: Artificial Sequence
300 <220> FEATURE:
301 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
303 <400> SEQUENCE: 23
304 Ala Leu Pro Ala Leu Pro Gln
305   1              5
307 <210> SEQ ID NO: 24
308 <211> LENGTH: 7
309 <212> TYPE: PRT
310 <213> ORGANISM: Artificial Sequence
312 <220> FEATURE:
313 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
315 <400> SEQUENCE: 24
316 Val Leu Pro Ala Ala Pro Gln
317   1              5
319 <210> SEQ ID NO: 25
320 <211> LENGTH: 7
321 <212> TYPE: PRT
322 <213> ORGANISM: Artificial Sequence
324 <220> FEATURE:
325 <223> OTHER INFORMATION: Description of Artificial Sequence: oligopeptide
327 <400> SEQUENCE: 25
328 Val Leu Pro Ala Leu Ala Gln
```

RAW SEQUENCE LISTING ERROR SUMMARY    DATE: 10/30/2002
PATENT APPLICATION:  US/10/028,075A    TIME: 15:01:57

Input Set : **A:\EP.txt**
Output Set: **N:\CRF4\10302002\J028075A.raw**

**Please Note:**
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:97; Xaa Pos. 2
Seq#:98; Xaa Pos. 2

**Invalid Line Length:**
The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:2; Line(s) 42
Seq#:103; Line(s) 1376
Seq#:118; Line(s) 1565

**VERIFICATION SUMMARY**                    DATE: 10/30/2002
PATENT APPLICATION:  **US/10/028,075A**          TIME: 15:01:57

Input Set : **A:\EP.txt**
Output Set: **N:\CRF4\10302002\J028075A.raw**

L:1287 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:97
L:1292 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:97 after pos.:0
L:1312 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:98 after pos.:0